

22

A dynamical-system picture of a simple
branching-process phase transitionDavid Williams^a**Abstract**

This paper proves certain results from the ‘appetizer for non-linear Wiener–Hopf theory’, [5]. Like that paper, it considers only the simplest possible case in which the underlying Markov process is a two-state Markov chain. Key generating functions provide solutions of a simple two-dimensional dynamical system, and the main interest is in the way in which Probability Theory and ODE theory complement each other. No knowledge of either ODE theory or Wiener–Hopf theory is assumed. Theorem 1.1 describes one aspect of a phase transition which is more strikingly conveyed by Figures 4.1 and 4.2.

AMS subject classification (MSC2010) 60J80, 34A34

1 Introduction

This paper is a development of something I mentioned briefly in talks I gave at Bristol, when John Kingman was in the audience, and at the Waves conference in honour of John Toland at Bath. I thanked both John K and John T for splendid mathematics and for their wisdom and kindness.

The main point of the paper is to prove Theorem 1.1 and related results in a way which emphasizes connections with a simple dynamical system. The phase transition between Figures 4.1 and 4.2 looks more dramatic than the famous 1-dimensional result we teach to all students.

^a Mathematics Department, Swansea University, Singleton Park, Swansea SA2 8PP; dw@reynoldston.com

The model studied here is a special case of the model introduced in [5]. I called that paper, which contained no proofs, an ‘appetizer’; but before writing a fuller version, I became caught up in Jonathan Warren’s enthusiasm for the relevance of *complex* dynamical systems (in \mathbb{C}^2). See , Warren and Williams [4]. This present paper, *completely independent of the earlier appetizer* and of my paper with Warren, can, I hope, provide a more tempting appetizer for what I called ‘non-linear Wiener–Hopf theory’. No knowledge of any kind of Wiener–Hopf theory is assumed here.

I hope that Simon Harris and I can throw further light on the models considered here, on the other models in [5], and on still other, quite different, models.

Our model. A particle moving on the real line can either be of type $+$ in which case it moves right at speed 1 or of type $-$ in which case it moves left at speed 1.

Let q_- and q_+ be fixed numbers with $q_- > q_+ > 0$, and let β be a positive parameter. We write $K_{\pm} = q_{\pm} + \beta$. So, to display things, we have

$$q_- > q_+ > 0, \quad \beta > 0, \quad K_+ = q_+ + \beta, \quad K_- = q_- + \beta. \quad (1.1)$$

We define

$$\beta_c := \frac{1}{2} (\sqrt{q_-} - \sqrt{q_+})^2.$$

A particle of type \pm can flip to the ‘opposite’ type at rate q_{\pm} and can, at rate β , die and at its death give birth to two daughter particles (of the same type and position as their ‘parent’). This is why β is a ‘birth rate’. The usual independence conditions hold.

Theorem 1.1 *Suppose that our process starts at time 0 with just 1 particle of type $+$ at position 0.*

- (a) *Suppose that $\beta > \beta_c$. Then, with probability 1, each of infinitely many particles will spend time to the left of 0.*
- (b) *Suppose instead that $\beta \leq \beta_c$. Then, with probability not less than $1 - \sqrt{q_+/q_-}$, there will never be any particles to the left of 0.*

Large-deviation theory (of which the only bit we need is proved here) allows one to prove easily that if $\beta < \beta_c$, then, almost surely, only a finite number of particles are ever to the left of 0.

The interplay between the Probability and the ODE theory is what is

most interesting. We shall see that β_c plays the rôle of a critical parameter in several ways, some probabilistic, some geometric. The ‘balance’ which occurs when $\beta = \beta_c$ is rather remarkable.

The paper poses a tantalizing problem which I cannot yet solve.

2 Wiener–Hopferization

2.1 The processes $\{N^\pm(\varphi) : \varphi \geq 0\}$

For any particle i alive at time t , we define $\Phi_i(t)$ to be its position on the real line at time t , and we extend the definition of Φ_i by saying that at any time s before that particle’s birth, $\Phi_i(s)$ is the position of its unique ancestor alive at time s .

So far, so sane! But we are now going to Wiener–Hopferize everything with a rather clumsy definition which defines for each $\varphi \geq 0$ two subsets, $S^+(\varphi, \beta)$ and $S^-(\varphi, \beta)$, of particles.

We put particle i in set $S^+(\varphi, \beta)$ if there is some t in $[B(i), D(i))$ where $B(i)$ and $D(i)$ are, respectively, the times of birth and death of particle i , such that

- $\Phi_i(t) = \varphi$,
- $\Phi_i(t) \geq \max\{\Phi_i(s) : s \leq t\}$, and
- Φ_i grows to the right of t in that, for $\varepsilon > 0$, there exists a δ with $0 < \delta < \varepsilon$ such that $\Phi_i(\cdot) > \varphi$ throughout $(t, t + \delta)$.

At the risk of labouring things, let me describe $S^-(\varphi, \beta)$ for $\varphi \geq 0$. We put particle i in set $S^-(\varphi, \beta)$ if there is some t in $[B(i), D(i))$ such that $\Phi_i(t) = -\varphi$, $\Phi_i(t) \leq \min\{\Phi_i(s) : s \leq t\}$, and Φ_i decreases to the right of t in that, for $\varepsilon > 0$, there exists a δ with $0 < \delta < \varepsilon$ such that $\Phi_i(\cdot) < -\varphi$ throughout $(t, t + \delta)$.

Of course, there may be particles not in $\bigcup_{\varphi \geq 0} \{S^+(\varphi, \beta) \cup S^-(\varphi, \beta)\}$.

We define $N^+(\varphi, \beta)$ [respectively, $N^-(\varphi, \beta)$] to be the number of particles in $S^+(\varphi, \beta)$ [resp., $S^-(\varphi, \beta)$].

We let \mathbb{P}_β^+ [respectively, \mathbb{P}_β^-] be the probability law of our model when it starts with 1 particle of type + [resp., –] at position 0 at time 0; and we let \mathbb{E}_β^+ [resp., \mathbb{E}_β^-] be the associated expectation.

We often suppress the ‘ β ’ in the notation for \mathbb{P}_β^\pm , \mathbb{E}_β^\pm , S_β^\pm , N_β^\pm .

Then, under \mathbb{P}^+ , $N^+ = \{N^+(\varphi) : \varphi \geq 0\}$ is a standard branching process, in which a particle dies at rate K_+ and is replaced at the ‘ Φ -time’ of its death by a random non-negative number, possibly 1 and possibly

∞ , of children, the numbers of children being independent, identically distributed random variables. I take this as intuitively obvious, and I am not going to ruin the paper by spelling out a proof.

Note that in the \mathbb{P}^- branching process $N^- = \{N^-(\varphi) : \varphi \geq 0\}$, a particle may die without giving birth.

For $0 \leq \theta < 1$, define

$$g^{++}(\varphi, \theta) := \mathbb{E}^+ \theta^{N^+(\varphi)}, \quad h^{-+}(\varphi, \theta) := \mathbb{E}^- \theta^{N^+(\varphi)}.$$

Clearly, for $0 \leq \theta < 1$,

$$\begin{aligned} h^{-+}(\varphi, \theta) &= \mathbb{E}^- \mathbb{E}^- [\theta^{N^+(\varphi)} \mid N^+(0)] = \mathbb{E}^- g^{++}(\varphi, \theta)^{N^+(0)} \\ &= H^{-+}(g^{++}(\varphi, \theta)), \end{aligned}$$

where

$$H^{-+}(\theta) = \mathbb{E}^- \theta^{N^+(0)} = \sum h_n^{-+} \theta^n,$$

where

$$h_n^{-+} := \mathbb{P}^- [N^+(0) = n].$$

It may well be that $h_\infty^{-+} := \mathbb{P}^- [N^+(0) = \infty] > 0$. Note that

$$H^{-+}(1-) := \lim_{\theta \uparrow 1} H^{-+}(\theta) = \mathbb{P}^- [N^+(0) < \infty].$$

2.2 The dynamical system

We now take θ in $(0, 1)$ and derive the backward differential equations for

$$x(\varphi) := g^{++}(\varphi, \theta), \quad y(\varphi) := h^{-+}(\varphi, \theta),$$

in the good old way in which we teach Applied Probability, and then study the equations. [5] looks a bit more ‘rigorous’ here.

Consideration of what happens between times 0 and dt tells us that

$$x(\varphi + d\varphi) = \{1 - K_+ d\varphi\}x(\varphi) + \{q_+ d\varphi\}y(\varphi) + \{\beta d\varphi\}x(\varphi)^2 + o(d\varphi).$$

The point here is of course that if we started with 2 particles in the $+$ state, then $\mathbb{E} \theta^{N^+(\varphi)} = x(\varphi)^2$. We see that, with x' meaning $x'(\varphi)$,

$$x' = q_+(y - x) + \beta(x^2 - x). \quad (2.1a)$$

Similarly, remembering that Φ starts to run backwards when the particle starts in state $-$, we find that

$$y(\varphi - d\varphi) = \{1 - K_- d\varphi\}y(\varphi) + \{q_- d\varphi\}x(\varphi) + \{\beta d\varphi\}y(\varphi)^2 + o(d\varphi),$$

whence

$$-y' = q_-(x - y) + \beta(y^2 - y). \quad (2.1b)$$

Of course, $y = H^{-+}(x)$ must represent the track of an integral curve of the *dynamical system* (2.1), and since $y' = H^{-+}'(x)x'$, we have an *autonomous equation for H^{-+}* which we shall utilize below.

Note that the symmetry of the situation shows that $x = H^{+-}(y)$ must also represent the track of an integral curve of our dynamical system, though one traversed in the ' φ -reversed' direction.

Probability Theory guarantees the existence of the 'probabilistic solutions' of the dynamical system tracking curves $y = H^{-+}(x)$ and $x = H^{+-}(y)$.

Lemma 2.1 *There can be no equilibria of our dynamical system in the interior of the unit square.*

Proof For if (x, y) is in the interior and

$$q_+(y - x) + \beta(x^2 - x) = 0, \quad q_-(x - y) + \beta(y^2 - y) = 0,$$

then $y \geq x$ from the first equation and $x \geq y$ from the second. Hence $x = y$ and $x^2 - x = y^2 - y = 0$. \square

2.3 Change of θ

We need to think about how a change of θ would affect things. Suppose that $\alpha = \mathbb{E}^+ \theta^{N^+(\psi)}$ where $0 < \alpha < 1$. Then

$$\mathbb{E}^+ \mathbb{E}^+ [\theta^{N^+(\varphi+\psi)} \mid N^+(\varphi)] = \mathbb{E}^+ \alpha^{N^+(\varphi)} = g^{++}(\varphi, \alpha),$$

where $\alpha = g^{++}(\psi, \theta)$. So, we have the *probabilistic-flow relation*

$$g^{++}(\varphi + \psi, \theta) = g^{++}(\varphi, g^{++}(\psi, \theta)). \quad (2.2)$$

Likewise, $h^{-+}(\varphi + \psi, \theta) = h^{-+}(\varphi, g^{++}(\psi, \theta))$. Thus, changing from θ to $\alpha = \mathbb{E}^+ \theta^{N^+(\psi)}$ just changes the starting-point of the motion along the probabilistic curve from $(\theta, H^{-+}(\theta))$ to the point $(\alpha, H^{-+}(\alpha))$ still on the probabilistic curve. This is why we may sometimes seem not to care about θ , and why it is not in our notation for $x(\varphi)$, $y(\varphi)$. But we shall discuss θ when necessary, and the extreme values 0 and 1 of θ in Subsection 4.6.

If for *any* starting point $\mathbf{v}_0 = (x_0, y_0)$ within the unit square, we write $\mathbf{V}(\varphi, \mathbf{v}_0)$ for the value of $(x(\varphi), y(\varphi))$, then, for values of φ and ψ

in which we are interested, we have (granted existence and uniqueness theorems) the *ODE-flow relation*

$$\mathbf{V}(\varphi + \psi, \mathbf{v}_0) = \mathbf{V}(\varphi, \mathbf{V}(\psi, \mathbf{v}_0))$$

which generalizes (2.2). (The possibility of explosions need not concern us: we are interested only in what happens within the unit square.) For background on ODE flows, see [1].

3 How does ODE theory see the phase transition?

3.1 An existence theorem

Even if you skip the (actually quite interesting!) proof of the following theorem, do not skip the discussion of the result which makes up the next subsection.

Theorem 3.1 *There exist constants $\{a_n : n \geq 0\}$ with $a_0 = 0$, all other a_n strictly positive, and*

$$\sum a_n \leq q_+/(q_- + \beta), \quad (3.1)$$

and a solution $(x(\varphi), y(\varphi))$ of the ‘ φ -reversed’ dynamical system

$$-x' = q_+(y - x) + \beta(x^2 - x), \quad y' = q_-(x - y) + \beta(y^2 - y),$$

such that $x(\varphi) = A(y(\varphi))$, where we now write $A(y) = \sum a_n y^n$.

Proof Assume that constants a_n as described exist. Since $x'(\varphi) = A'(y(\varphi))y'(\varphi)$, we have

$$-\{q_+y + \beta A(y)^2 - K_+A(y)\} = A'(y) \{q_-A(y) + \beta y^2 - K_-y\}.$$

Comparing coefficients of y^0 ,

$$\beta a_0^2 - K_+a_0 = -a_1q_-a_0,$$

and we are guaranteeing this by taking $a_0 = 0$. Comparing coefficients of y^1 , we obtain

$$q_-a_1^2 - (K_- + K_+)a_1 + q_+ = 0.$$

We take

$$\begin{aligned} a_1 &= \frac{K_- + K_+ - \sqrt{(K_- + K_+)^2 - 4q_-q_+}}{2q_-} \\ &= \frac{2q_+}{K_- + K_+ + \sqrt{(K_- + K_+)^2 - 4q_-q_+}} \end{aligned}$$

from which it is obvious that $0 < a_1 < 1$.

On comparing coefficients of y^n we find that, for $n \geq 2$,

$$\begin{aligned} & \{K_+ + nK_- - (n+1)q_- a_1\} a_n \\ &= \beta \sum_{k=1}^{n-1} a_k a_{n-k} + q_- \sum_{k=1}^{n-2} (k+1) a_{k+1} a_{n-k} + \beta(n-1) a_{n-1}. \end{aligned}$$

We now consider the a_n as being *defined* by these recurrence relations (and the values of a_0 and a_1). It is clear that the a_n are all positive.

Temporarily fix $N > 2$, and define

$$\begin{aligned} A_N(y) &:= \sum_{n=0}^N a_n y^n, \\ L(y) &:= \sum_{n=0}^N \ell_n y^n := -q_+ y - \beta A_N(y)^2 + K_+ A_N(y), \\ R(y) &:= \sum_{n=0}^N r_n y^n := A'_N(y) \{q_- A_N(y) + \beta y^2 - K_- y\}. \end{aligned}$$

For $n \leq N$ we have $\ell_n = r_n$ by the recurrence relations. It is clear that for $n > N$ we have $\ell_n \leq 0$ and $r_n \geq 0$. Hence, for all y in $(0, 1)$,

$$-q_+ y - \beta A_N(y)^2 + K_+ A_N(y) \leq A'_N(y) \{q_- A_N(y) + \beta y^2 - K_- y\}. \quad (3.2)$$

Suppose for the purpose of contradiction that there exists y_0 in $(0, 1)$ with $A_N(y_0) = y_0$. Then

$$-q_+ - \beta y_0 + K_+ \leq A'_N(y_0) \{q_- + \beta y_0 - K_-\}.$$

However, the left-hand side is positive while the right-hand side is negative.

Because $A_N(0) = 0$ and $A'_N(0) = a_1 < 1$, the contradiction establishes that $A_N(y) < y$ for $y \in (0, 1)$, so that $A_N(1) \leq 1$. Since this is true for every N , and each a_n ($n > 1$) is strictly positive, we have $A_N(1) < 1$ for every N .

By inequality (3.2), we have

$$D_N q_- (A_N - 1) + q_+ + \beta A_N^2 - K_+ A_N \geq 0,$$

where $A_N := A_N(1) < 1$ and $D_N := A'_N(1)$. Because each a_n ($n > 0$) is positive it is clear that $A_N < D_N$. We therefore have

$$q_+ + \beta A_N^2 - K_+ A_N \geq D_N q_- (1 - A_N) \geq q_- A_N (1 - A_N),$$

which simplifies to

$$(1 - A_N) \{q_+ - (\beta + q_-) A_N\} \geq 0.$$

Since $(1 - A_N) > 0$, we have $(\beta + q_-)A_N \leq q_+$, and result (3.1) follows.

It is clear that we now need to consider the autonomous equation for $y = y(\varphi)$:

$$y' = q_-[A(y) - y] + \beta[y^2 - y], \quad y(0) = \theta.$$

But we can describe $y(\varphi)$ as $\mathbb{E}\theta^{Z(\varphi)}$ where $\{Z(\varphi) : \varphi \geq 0\}$ is a classical branching process in which (with the usual independence properties) a particle dies at rate K_- and at the moment of its death gives birth to C children where

$$\mathbb{P}(C = n) = \begin{cases} q_- a_n / K_- & \text{if } 1 \leq n \leq \infty \text{ and } n \neq 2, \\ (\beta + q_- a_2) / K_- & \text{if } n = 2. \end{cases}$$

Of course, $a_\infty = 1 - \sum a_n$.

Then $(x(\varphi), y(\varphi)) = (A(y(\varphi)), y(\varphi))$ describes the desired solution starting from $(A(\theta), \theta)$. \square

3.2 Important discussion

Of course, ODE theory cannot see what we shall see later: namely that $A(\cdot) = H^{+-}(\cdot)$ when $\beta > \beta_c$ but $A(\cdot) \neq H^{+-}(\cdot)$ when $\beta \leq \beta_c$. When $\beta > \beta_c$, the curve $x = H^{+-}(y)$ is the steep bold curve $x = A(y)$ at the *left*-hand side of the picture as in Figure 4.1. But when $\beta \leq \beta_c$, the curve $x = H^{+-}(y)$ is the steep bold curve at the *right*-hand side of the picture as in Figure 4.2. Ignore the shaded triangle for now.

What ODE theory must see is that whereas there is only one integral curve linking the top and bottom of the unit square when $\beta > \beta_c$, there are infinitely many such curves when $\beta \leq \beta_c$ of which two, the curves $x = A(y)$ and $x = H^{+-}(y)$, derive from probability generating functions (pgfs).

It does not seem at all easy to prove by Analysis that, when $\beta \leq \beta_c$, there is an integral curve linking the bottom of the unit square to the point $(1, 1)$, of the form $x = F(y)$ where F is the pgf of a random variable which can perhaps take the value ∞ . Methods such as that used to prove Theorem 3.1 will not work.

Moreover, it is not easy to compute $H^{+-}(0)$ when β is equal to, or close to, β_c . If for example, $q_+ = 1$, $q_- = 4$ and $\beta = 0.4$, then one can be certain that $H^{+-}(0) = 0.6182$ to 4 places, and indeed one can easily calculate it to arbitrary accuracy. But the critical nature of β_c shows itself in unstable behaviour of some naïve computer programs when β is

equal to, or close to, β_c . I believe that in the critical case when $q_+ = 1$, $q_- = 4$ and $\beta = 0.5$, $H^{+-}(0)$ is just above 0.6290.

Mathematica is understandably extremely cautious in regard to the non-linear dynamical system (2.1), and drives one crazy with warnings. If forced to produce pictures, it can produce some rather crazy ones, though usually, but not absolutely always, under protest. Its pictures can be coaxed to agree with those in the earlier appetizer which were produced from my own ‘C’ Runge–Kutta program which yielded Postscript output. Sadly, that program and lots of others were lost in a computer burn-out before I backed them up.

4 Proof of Theorem 1.1 and more

4.1 When $\beta > \beta_c$

Lemma 4.1 When $\beta > \beta_c$,

- (a) $H^{-+}(0) = \mathbb{P}^{-}[N^{+}(0) = 0] = 0$,
- (b) $H^{-+}(1-) = \mathbb{P}^{-}[N^{+}(0) < \infty] < 1$,
- (c) $H^{+-}(1-) = \mathbb{P}^{+}[N^{-}(0) < \infty] < 1$,
- (d) $H^{+-}(0) = \mathbb{P}^{+}[N^{-}(0) = 0] = 0$.

It is clearly enough to prove the lemma under the assumption

$$\frac{1}{2}(\sqrt{q_-} - \sqrt{q_+})^2 < \beta < \frac{1}{2}(\sqrt{q_-} + \sqrt{q_+})^2,$$

and this is made throughout the proof.

Proof Result (a) is obvious.

The point $(1, 1)$ is an equilibrium point of our dynamical system, and we consider the linearization of the system near this equilibrium. We put $x = 1 + \xi$, $y = 1 + \eta$ and linearize by ignoring terms in ξ^2 and η^2 :

$$\begin{pmatrix} \xi' \\ \eta' \end{pmatrix} = \begin{pmatrix} -q_+ + \beta & q_+ \\ -q_- & q_- - \beta \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix},$$

the matrix being the linearization matrix of our system at $(1, 1)$. The characteristic equation for the eigenvalues of this matrix is

$$\lambda^2 + (q_+ - q_-)\lambda + (q_- + q_+)\beta - \beta^2 = 0.$$

The discriminant ‘ $B^2 - 4AC$ ’ is

$$\{2\beta - (q_+ + q_-)\}^2 - 4q_-q_+.$$

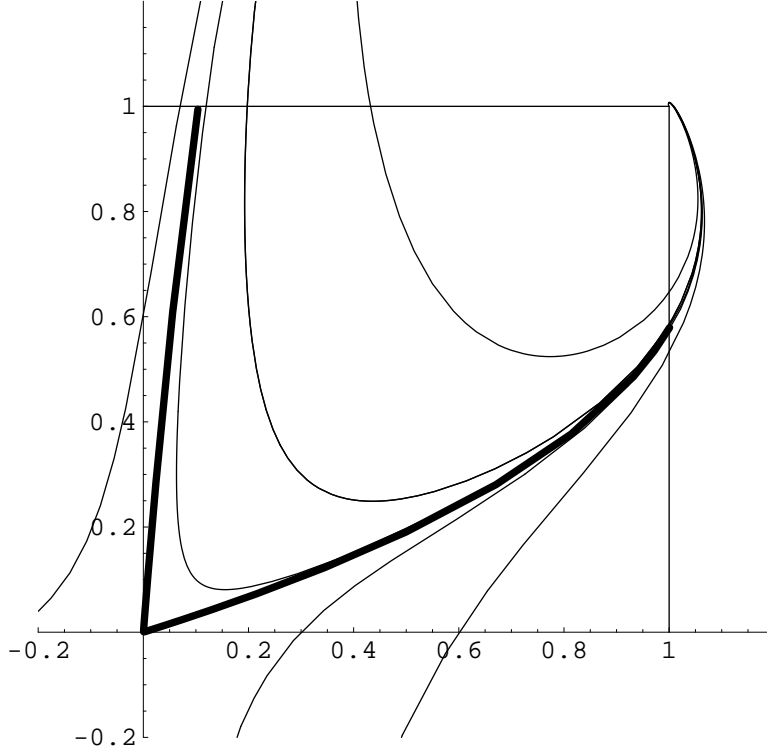


Figure 4.1 *Mathematica* picture of the supercritical case when $q_+ = 1$, $q_- = 4$, $\beta = 4$

This expression is zero if $\beta = \frac{1}{2}(\sqrt{q_-} \pm \sqrt{q_+})^2$. So, in our case, the eigenvalues λ have non-zero imaginary parts. Any solution of our system converging to $(1, 1)$ as $\varphi \rightarrow \pm\infty$ must spiral, and cannot remain inside the unit square. Hence results (b) and (c) hold.

It is now topologically obvious (since there are no equilibria inside the unit square) that we must have $\mathbb{P}^+[N^-(0) = 0] = 0$; otherwise how could the curve $x = H^{+-}(y)$ link the top and bottom edges of the square? Thus result (d) holds. \square

Of course, we can now deduce from result (3.1) that (when $\beta > \beta_c$)

$$H^{+-}(1-) = \mathbb{P}^+[N^-(0) < \infty] \leq q_+/(q_- + \beta).$$

Figure 4.1, which required ‘cooking’ beyond choosing different φ -ranges for different curves, represents the case when $q_+ = 1$, $q_- = 4$

and $\beta = 4$. The lower bold curve represents $y = H^{-+}(x)$ and the upper $x = H^{+-}(y)$. As mentioned previously, $H^{+-}(y) = A(y)$ where A is the function of Theorem 3.1.

The motion along the lower probabilistic curve $y = H^{-+}(x)$ will start at $(\theta, H^{-+}(\theta))$ and move towards $(0, 0)$ converging to $(0, 0)$ as $\varphi \rightarrow \infty$ since $N^\pm(\varphi) \rightarrow \infty$. If we fix φ and let $\theta \rightarrow 1$, we move along the curve towards the point $(1, H(1-))$. Of course, we could alternatively leave θ fixed and run φ backwards. It is clear because of the spiralling around $(1, 1)$ that the power series $H^{-+}(x)$ must have a singularity at some point x not far to the right of 1.

Motion of the dynamical system along the steep probabilistic curve $x = H^{+-}(y)$ on the left of the picture will be upwards because it is the φ -reversal of the natural probabilistic motion. Now you understand the sweep of the curves in the top-right of the picture.

4.2 A simple large-deviation result

Let $\{X(t) : t \geq 0\}$ be a Markov chain on $\{+, -\}$ with Q -matrix

$$Q = \begin{pmatrix} -q_+ & q_+ \\ q_- & -q_- \end{pmatrix}.$$

Let V be the function on $\{+, -\}$ with $V(+) = 1$ and $V(-) = -1$ and define $\Phi_X(t) = \int_0^t V(X_s) ds$. Almost surely, $\Phi_X(t) \rightarrow \infty$. We stay in ‘dynamical-system mode’ to obtain the appropriate Feynman–Kac formula.

Let $\mu > 0$, and define (with the obvious meanings of \mathbb{E}^\pm)

$$u(t) = \mathbb{E}^+ \exp\{-\mu \Phi_X(t)\}, \quad v(t) = \mathbb{E}^- \exp\{-\mu \Phi_X(t)\}.$$

Then

$$u(t + dt) = \{1 - q_+ dt\} e^{-\mu dt} u(t) + q_+ dt v(t) + o(dt),$$

with a similar equation for v . We find that

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} -q_+ - \mu & q_+ \\ q_- & -q_- + \mu \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}, \quad \begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \exp\{t(Q - \mu V)\} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

where V also denotes the operator $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ of multiplication by the function V .

Lemma 4.2 *If $\beta < \beta_c := \frac{1}{2} (\sqrt{q_-} - \sqrt{q_+})^2$, then there exist positive*

constants ε, κ, A such that

$$e^{\beta t} \mathbb{P}^\pm[\Phi_X(t) \leq \varepsilon t] \leq Ae^{-\kappa t}.$$

Proof We have just shown that

$$\mathbb{E}^\pm[e^{-\mu\Phi_X(t)} f(X_t)] = \exp\{t(Q - \mu V)\} f.$$

Now $Q - \mu V$ has larger eigenvalue

$$\gamma = -\frac{1}{2}(q_- + q_+) + \frac{1}{2}\sqrt{(q_- + q_+)^2 - 4(q_- - q_+)\mu + 4\mu^2}.$$

We fix μ at $\frac{1}{2}(q_- - q_+)$ to obtain the minimum value $\frac{1}{2}(\sqrt{q_-} - \sqrt{q_+})^2$ of γ . Hence, for $\varepsilon > 0$ and some constant A_ε ,

$$\begin{aligned} \mathbb{P}^\pm[\Phi_X(t) \leq \varepsilon t] &= \mathbb{P}^\pm[\mu(\varepsilon t - \Phi_X(t)) \geq 0] \leq \mathbb{E}^\pm \exp\{\mu\varepsilon t - \mu\Phi_X(t)\} \\ &\leq A_\varepsilon \exp\left\{\frac{1}{2}\varepsilon(q_- - q_+)t - \frac{1}{2}(\sqrt{q_-} - \sqrt{q_+})^2 t\right\}. \end{aligned}$$

The lemma follows. \square

For a fine paper proving very precise large-deviation results for Markov chains via explicit calculation, see Brydges, van der Hofstad, and König [2].

4.3 When $\beta < \beta_c$

Lemma 4.3 When $\beta \leq \beta_c$,

- (a) $H^{-+}(0) = \mathbb{P}^-[N^+(0) = 0] = 0$,
- (b) $H^{-+}(1-) = \mathbb{P}^-[N^+(0) < \infty] = 1$,
- (c) $H^{+-}(0) = \mathbb{P}^+[N^-(0) = 0] > 0$,
- (d) $H^{+-}(1-) = \mathbb{P}^+[N^-(0) < \infty] = 1$.

Note Though Figure 4.2 relates to the case when $\beta = \beta_c$, pictures for the subcritical case when $\beta < \beta_c$ look very much the same.

Proof Result (a) remains trivial.

By Lemma 4.2, there exist $\varepsilon > 0, \kappa > 0$ and $A > 0$ such that, for a single particle moving according to Q -matrix Q , we have

$$e^{\beta t} \mathbb{P}[\Phi_X(t) \leq \varepsilon t] \leq Ae^{-\kappa t}.$$

For the branching process, the expression on the left-hand side is the expected number of particles with Φ -value less than or equal to εt at real time t . So the probability that some particle has Φ -value less than or equal to t is at most $Ae^{-\kappa t}$.

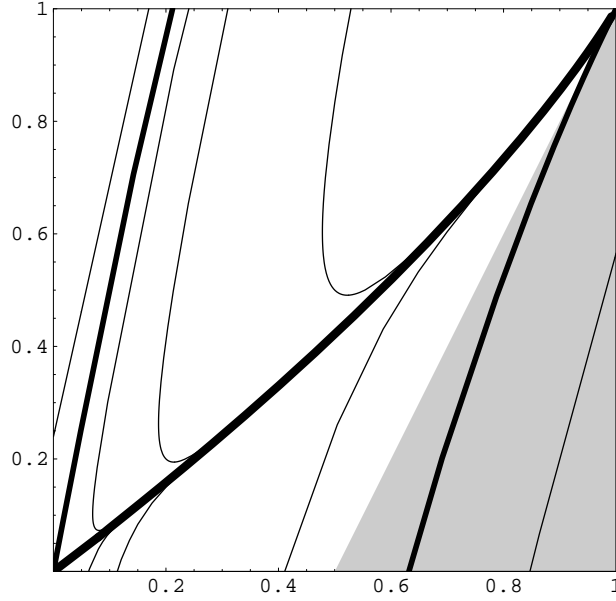


Figure 4.2 *Mathematica* picture of the critical case when $q_+ = 1$, $q_- = 4$, $\beta = 0.5$

By the Borel–Cantelli Lemma, there will almost surely be a random positive integer n_0 such that for all $n \geq n_0$, every particle alive at real time n will have Φ -value greater than εn . Since Φ can only move left at speed 1, there must almost surely come a time after which no particle has a positive Φ -value. Hence $\mathbb{P}^-[N^+(0) = \infty] = 0$, and result (b) is proved.

Now suppose for the purpose of contradiction that $\mathbb{P}^+[N^-(0) > 0] = 1$. Since a particle started at state $+$ can remain there for an arbitrary long time without giving birth, it follows that any particle in the $+$ state and with any positive Φ -value will have a descendant for which Φ will become negative. This contradicts what we proved in the previous paragraph, so result (c) is established.

Since the $y = H^{-+}(x)$ curve connects $(1, 1)$ to $(0, 0)$ and the other probabilistic curve $x = H^{+-}(y)$ starts at $(H^{+-}(0), 0)$ where $H^{+-}(0) > 0$, and since these curves cannot cross at an interior point of the unit square, it must be the case that $H^{+-}(1-) = 1$, so that property (d) holds. \square

In the analogue of Figure 4.2 for a subcritical case (which, as I have said, looks very much like Figure 4.2), motion along the higher probabilistic curve $y = H^{-+}(x)$ will again start at $(\theta, H^{-+}(\theta))$ and move towards $(0, 0)$, this because $N^+(\varphi) \rightarrow \infty$. Since $N^-(\varphi) \rightarrow 0$, the natural probabilistic motion of the lower curve $x = H^{+-}(y)$ will converge to $(1, 1)$; but the φ -reversal means that the dynamical system will move downwards along this curve.

Sketch of geometric proof that $H^{-+}(1-) = 1$ if $\beta \leq \beta_c$ It is enough to prove the result when $\beta = \beta_c$. Let $m = \sqrt{q_-/q_+}$, the slope of the unique eigenvector of the linearity matrix at $(1, 1)$. Draw the line of slope m from $(1, 1)$ down to the y -axis, the sloping side of the shaded triangle in the picture. Now it is particularly easy to check that at any point of the sloping side the (dy/dx) -slope of an integral curve is greater than m . If the convex curve $y = H^{-+}(x)$ hit the vertical side of the triangle at any point lower than 1, we would have ‘contradiction of slopes’ where it crossed that sloping side. \square

4.4 Nested models and continuity at phase transition

Take $\beta_0 > \beta_c$ and let M_{β_0} be our model with initial law $\mathbb{P}_{\beta_0}^+$ (in the obvious sense). Label birth-times T_1, T_2, T_3, \dots in the order in which they occur, and for each n call one of the two children born at T_n ‘first’, the other ‘second’. Let U_1, U_2, U_3, \dots be independent random variables each with the uniform distribution on $[0, 1)$. We construct a nested family of models $\{M_\beta : \beta \leq \beta_0\}$ as follows.

Fix $\beta < \beta_0$ for the moment. If $U_n > \beta/\beta_0$, erase the whole family tree in M_{β_0} descended from the second child born at time T_n . Of course, this family tree may already have been erased earlier. In this way, we have a model M_β with desired law \mathbb{P}_β^+ . The set $S^-(\varphi, \beta)$ will now denote the $S^-(\varphi)$ set for the ‘nested’ model M_β , and $N^-(\varphi, \beta)$ will denote its cardinality.

A particle i contributing to $S^-(0, \beta)$ determines a path in $\{+, -\} \times [0, \infty)$:

$$\{(\text{Ancestor}_i(t), \Phi_i(t)) : t < \rho_i\}$$

where ρ_i is the first time after which Φ_i becomes negative. Along that M_β -path, there will be finitely many births. Now, for fixed β it is almost surely true that $U_n \neq \beta$ for all n . It is therefore clear that, almost surely, for $\beta' < \beta$ and β' sufficiently close to β , the $M_{\beta'}$ -path will also be a path

of $M_{\beta'}$. In other words, we have the left-continuity property

$$S^-(0, \beta) = \bigcup_{\beta' < \beta} S^-(0, \beta'), \text{ almost surely.}$$

It therefore follows from the Monotone-Convergence Theorem that

$$\mathbb{E}^+ N^-(0, \beta_c) = \uparrow \lim_{\beta \uparrow \beta_c} \mathbb{E}^+ N^-(0, \beta). \quad (4.1)$$

Clearly, something goes seriously wrong in regard to right-continuity at β_c . Suppose we have a path which contributes to $S^-(0, \beta)$ for all $\beta > \beta_c$. Then, for all birth-times T_n along that path we have $U_n \leq \beta/\beta_0$ for all $\beta > \beta_c$ and hence $U_n \leq \beta_c/\beta_0$. Hence

$$S^-(0, \beta_c) = \bigcap_{\beta > \beta_c} S^-(0, \beta).$$

But it is clearly possible to have a decreasing sequence of infinite sets with finite intersection. And recall that (more generally) the Monotone-Convergence Theorem is guaranteed to work ‘downwards’ (via the Dominated-Convergence Theorem) only when one of the random variables has finite expectation.

4.5 Expectations and an embedded discrete-parameter branching process

If either of the curves $y = H^{-+}(x)$ or $x = H^{+-}(y)$ approaches $(1, 1)$, it must do so in a definite direction and it is well known (and an immediate consequence of l’Hôpital’s rule) that that direction must be an eigenvector of the linearity matrix at $(1, 1)$. When $\beta = \beta_c$, there is (as we have seen before) only one eigenvector $\begin{pmatrix} m \\ 1 \end{pmatrix}$ with $m = \sqrt{q_-/q_+}$. Thus

$$\mathbb{E}^- N^+(0, \beta_c) = (H^{-+})'(1, \beta_c) = m = \sqrt{q_-/q_+}. \quad (4.2)$$

We know that if $\beta < \beta_c$, then $H^{+-}(1-) = 1$ and we can easily check that (as geometry would lead us to guess)

$$\mathbb{E}^+ N^-(0, \beta) = (H^{+-})'(1, \beta) \leq 1/m = \sqrt{q_+/q_-},$$

and now, by equation (4.1) we see that

$$\mathbb{E}^+ N^-(0, \beta_c) \leq 1/m = \sqrt{q_+/q_-}. \quad (4.3)$$

In particular, $\mathbb{P}^+[N^-(0, \beta_c) = \infty] = 0$, and so, in fact, $H^{+-}(1) = 1$ and we have equality at (4.3), whence

$$\mathbb{P}^+[N^-(0, \beta_c) \geq 1] \leq \mathbb{E}^+ N^-(0, \beta_c) = \sqrt{q_-/q_+}, \quad (4.4)$$

part of Theorem 1.1.

Now, for $\varphi > 0$, let

$$b(\varphi) = \mathbb{E}^+ N^+(\varphi), \quad c(\varphi) = \mathbb{E}^- N^+(\varphi).$$

Then

$$b'(\varphi) = q_+ \{c(\varphi) - b(\varphi)\} + \beta b(\varphi), \quad \text{etc.},$$

so that the linearization matrix at $(1, 1)$ controls expectations. We easily deduce the following theorem.

Theorem 4.4 *When $\beta = \beta_c$,*

$$\begin{aligned} \mathbb{E}^+ N^+(\varphi) &= e^{\frac{1}{2}(q_- - q_+)\varphi}, & \mathbb{E}^- N^+(\varphi) &= \sqrt{\frac{q_-}{q_+}} e^{\frac{1}{2}(q_- - q_+)\varphi}, \\ \mathbb{E}^+ N^-(\varphi) &= \sqrt{\frac{q_+}{q_-}} e^{-\frac{1}{2}(q_- - q_+)\varphi}, & \mathbb{E}^- N^-(\varphi) &= e^{-\frac{1}{2}(q_- - q_+)\varphi}. \end{aligned}$$

For any β , we can define the discrete-parameter branching-processes $\{W^\pm(n) : n \geq 0\}$ as follows. Let B_i be the birth time and D_i the death time of particle i . Recall that Φ_i is defined on $[0, D_i)$. Let $\sigma_i(0) = 0$ and, for $n \geq 1$, define

$$\sigma_i(n) := \inf\{t : B_i \leq t < D_i : t > \sigma_i(n-1); (-1)^n \Phi_i(t) > 0\},$$

with the usual convention that the infimum of the empty set is ∞ . Let

$$W^+(n) := \sharp\{i : \sigma_i(2n) < \infty\}, \quad W^-(n) := \sharp\{i : \sigma_i(2n+1) < \infty\}.$$

The ‘ W ’ notation is suggested by ‘winding operators’ in linear Wiener–Hopf theory.

Theorem 4.5 *W^\pm is a classical branching process under \mathbb{P}_β^+ , and is critical when $\beta = \beta_c$.*

The proof (left to the reader) obviously hinges on the case when $\varphi = 0$ of Theorem 4.4. And do have a think about the consequences of the ‘balance’

$$\mathbb{E}^- N^+(\varphi, \beta_c) \mathbb{E}^+ N^-(\varphi, \beta_c) = 1$$

in that theorem.

4.6 When $\theta = 0$ or 1

If we take $\theta = 0$ and set

$$b(\varphi) := \mathbb{P}^+[N^-(\varphi) = 0], \quad c(\varphi) := \mathbb{P}^-[N^-(\varphi) = 0],$$

then $\{(b(\varphi), c(\varphi)) : \varphi \geq 0\}$ is a solution of the φ -reversed dynamical system such that

$$b(\varphi) = H^{+-}(c(\varphi)).$$

When $\beta > \beta_c$, this solution stays at equilibrium point $(0, 0)$. When $\beta \leq \beta_c$, $(b(\varphi), c(\varphi))$ moves (as $0 \leq \varphi \uparrow \infty$) from $(H^{+-}(0), 0)$ to $(1, 1)$ tracing out the right-hand bold curve in the appropriate version of Figure 4.2.

When $\theta = 1$, $\{(B(\varphi), C(\varphi)) : \varphi \geq 0\}$, where

$$B(\varphi) := \mathbb{P}^+[N^-(\varphi) < \infty], \quad C(\varphi) := \mathbb{P}^-[N^-(\varphi) < \infty],$$

gives a solution of the φ -reversed dynamical system. When $\beta \leq \beta_c$, this solution stays at the equilibrium point $(1, 1)$. When $\beta > \beta_c$, $(B(\varphi), C(\varphi))$ moves (as $0 < \varphi \uparrow \infty$) from $(H^{+-}(1-), 1)$ to $(0, 0)$ tracing out the bold upper curve in the appropriate version of Figure 4.1. Of course, there is an appropriate version ('with + and - interchanged') for the lower curve.

4.7 A tantalizing question

When $\beta > \beta_c$, we have for the function $A(\cdot) = A(\cdot, \beta)$ of Theorem 3.1:

$$A(\theta, \beta) = \mathbb{E}^+ \theta^{N^-(0, \beta)}.$$

When $\beta \leq \beta_c$, we have, for $0 < \theta < 1$, $A(\theta, \beta) = \mathbb{E}^+ \theta^{Y_\beta}$ for some random variable Y_β . Can we find such a Y_β which is naturally related to our model? In particular, can we do this when $\beta = \beta_c$? It would be very illuminating if we could.

What is true for all $\beta > 0$ is that if X and Φ_X are as at the start of Subsection 4.2 and $\tau_X^-(0) := \inf\{t : \Phi_X(t) < 0\}$, then

$$\mathbb{E}^+ \exp\{-\beta \tau_X^-(0)\} = a_1,$$

with a_1 as in Theorem 3.1; and this tallies with $a_1 = \mathbb{P}^+[N^-(0) = 1]$ when $\beta > \beta_c$. Proof of the statements in this paragraph is left as an exercise.

Acknowledgements I certainly must thank a referee for pointing out some typos and, more importantly, a piece of craziness in my original version of a key definition. The referee also wished to draw our attention

(mine and yours) to an important survey article on tree-indexed processes [3] by Robin Pemantle. The process we have studied is, of course, built on the tree associated with the underlying branching process. But our Wiener–Hopferization makes for a rather unorthodox problem.

I repeat my thanks to Chris Rogers and John Toland for help with the previous appetizer. My thanks to Ian Davies and Ben Farrington for helpful discussions.

And I must thank the Henschel String Quartet — and not only for wonderful performances of Beethoven, Haydn, Mendelssohn, Schulhoff and Ravel, and fascinating discussions on music and mathematics generally and on the connections between the Mendelssohns and Dirichlet, Kummer, Hensel, Hayman. If I hadn’t witnessed something of the quartet’s astonishing dedication to music, it is very likely that I would have been content to leave things with that early appetizer and not made even the small advance which this paper represents. So, Happy music making, Henschels!

References

- [1] Arnol’d, V. I. 2006. *Ordinary Differential Equations*, transl. from Russian by R. Cooke. Universitext. Berlin: Springer-Verlag.
- [2] Brydges, D., Hofstad, R. van der, and König, W. 2007. Joint density for the local times of continuous-time Markov chains. *Ann. Probab.*, **35**, 1307–1332.
- [3] Pemantle, R. 1995. Tree-indexed processes. *Statist. Sci.*, **5**, 200–213.
- [4] Warren, J., and Williams, D. 2000. Probabilistic study of a dynamical system. *Proc. Lond. Math. Soc. (3)*, **81**(3), 618–650.
- [5] Williams, D. 1995. Non-linear Wiener–Hopf theory, 1: an appetizer. Pages 155–161 of: Azema, J., Émery, M., Meyer, P.-A., Yor, M. (eds), *Séminaire de probabilités* **29**. Lecture Notes in Math. **1613**. New York: Springer-Verlag.